Positional Cloning

The identification of a gene based solely on its position in the genome

Positional Cloning

Strength

No knowledge of the function of the gene product is required

Biochemistry, physiology, pathology, and biology all unnecessary

Positional Cloning

Weakness

No knowledge of the function of the gene product necessarily emerges

Insights regarding gene function rely on sequence similarity to gene products of known function

Genome Project continuously reduces the chances of this happening

What positional cloning can find

Mendelian disorders

Cystic Fibrosis, Neurofibromatosis, Huntington Disease, Duchenne Muscular Dystrophy, Hemochromatosis, Polyposis Coli, Polycystic Kidney Disease

Over 100 single-gene disorders, including all major genetic diseases in Caucasians and many of the major genetic diseases in non-Caucasian populations

Remaining Mendelian targets are all rare disorders

What positional cloning can find

Intermediate disorders

Mendelian versions of commonly non-Mendelian disorders

Alzheimer's in Volga Germans, MODY

Single gene disorders with non-Mendelian presentation

Torsion Dystonia, Hirschprung Disease

What positional cloning can find

Complex disorders

Mixtures of genetic and non-genetic causes

Non-Mendelian transmission - Alzheimer's, asthma, osteoporosis, psychiatric diseases, obesity, hypertension, prostate cancer, macular degeneration

Requirements for Positional Cloning

1. Evidence that a gene or genes are involved

Genetic epidemiology- Twin studies, adoption studies, segregation analysis

Risk ratio's for relatives - λ values

$$\lambda = \frac{\text{risk to a relative of an affected individual}}{\text{risk in the general population}}$$

λ_S values:	**Prostate Cancer**	**5**
	Schizophrenia	**9**
	Diabetes	**15**
	Autism	**75**
	Cystic Fibrosis	**400**
	Huntington's	**>1000**

Requirements for Positional Cloning

2. Families

How Many?

For Mendelian Traits

Each phase-known, fully informative, non-recombinant meiosis adds approximately + 0.3 to the LOD score

Since a LOD score of 3 is “proof” of linkage, 10 meioses will suffice under ideal conditions

Heterogeneity, recombination events, uninformative matings, and a number of other factors increase the number of individuals that need to be sampled

Requirements for Positional Cloning

2. Families

How Many?

Non-Mendelian Traits

Signal-to-noise problem

Affected siblings, affected pedigree members, many small families

Robust study designs typically incorporate 10^2 - 10^3 individuals

Positional Cloning

The Big Shortcut

Cytogenetic rearrangements

Translocations, Deletions

High resolution karyotyping

Syndromic presentation

The Positional Cloning Process

Step 1: Finding linkage

or

The first 3 orders of magnitude

Genotyping

Genetic markers - STPRP's

CA-repeats

Tetranucleotide repeats

Parameters of typical genome search - C.I.D.R., Marshfield

Linkage analysis - computerized evaluation

LIPED, LINKAGE, CRIMAP, SIBPAL, TDT, GENEHUNTER

Lod scores

acceptance of lod score of 3

good for finding linkage, terrible for finding genes

The Positional Cloning Process

Step 2: Narrowing the interval

or

The next order of magnitude

Typical linkage in disease families runs out of gas at ~ 1 Mb (30 genes)

Most common error is evaluating individual genes too soon

The Positional Cloning Process

Step 2: Narrowing the interval

Getting more meioses

Consanguineous families

Genetically isolated populations

Heterozygote advantage and linkage disequilibrium

Consanguineous families

Recessive disorders

Homozygosity mapping

Pooling strategies

Genetically isolated populations

Geographically isolated: Island populations
Mountain populations

Culturally isolated: Tribal populations
Bedouins
Native Americans
French Canadians

Religiously isolated: Iraqi Jews
Amish
Hutterites
Mennonites

Issues: effective founder size
time in isolation
relevance of that gene to disease in the general population

Heterozygote advantage and linkage disequilibrium

Can operate in very large, outbred populations

Hemochromatosis, cystic fibrosis, sickle cell disease

Final Decision on the Minimal Interval

Narrowing the interval

vs.

confidence in 2 individuals

Step 3: Finding all the genes

Build a physical contig to obtain entire region in cloned form

YAC's, P1's and PAC's, BAC's

Determine physical distances

Direct selection, exon trapping, and other clone-based methods

Large-scale sequencing plus informatics

Low-redundancy shotgun

BLAST searches for EST hits

EST's to complete genes

N.I.S.C.

Step 4: Identifying the responsible gene

DNA sequencing

cDNA vs. genomic

Association studies in unrelated affected individuals

Other Variations

Functional evaluation of candidate genes

Niemann-Pick Disease

Clues from model organisms

The Future

SNP's

Definition

Rationale

Numbers

Disease population association studies

Reliance on linkage disequilibrium - caveats

cSNP's